The Fundamental Idea of the Bipolar Junction Transistor

by

Nirmalya Basu

(nirmalya1basu@gmail.com)

This article describes the fundamental idea of the bipolar junction transistor (BJT) [1]. What is a bipolar junction transistor? It is a voltage-controlled current source. A device is called a "voltage-controlled current source" if the current flowing through a branch in it depends not on the voltage across that branch, but on the voltage between any two nodes in it which are not the end points of that branch. Now, how can we make it? Let us first figure out how we can make a current source, and then we can figure out how we can make it "voltage-controlled".

How can we make a current source? By using a device such that the current flowing through it is independent of the voltage across it. Now, what are the circuit elements that we learn about before we learn bipolar junction transistors? We learn about the resistor, the capacitor, the inductor and the diode. Now, we cannot use the resistor to make a current source. Why? Because its current-voltage relationship is $I = V/R$. This means that the current through a resistor depends on the voltage across it. So also for the capacitor and the inductor. This is because the current-voltage relationship of a capacitor is $I = C(dV/dt)$ and that of the inductor is $V = L(dI/dt)$. Now, let us come to the diode. The current-voltage relationship of a diode is

$$I = I_s \left(e^{\frac{V}{NV_T}} - 1 \right) \dots (1),$$

where I_s is a constant for a diode called the reverse saturation current of the diode,

V_T is a constant called the thermal voltage, and

N is a constant for a diode called the non-ideality factor of the diode.

Now, we know that the plot of this equation on the I-V plane is as follows (Fig. 1):

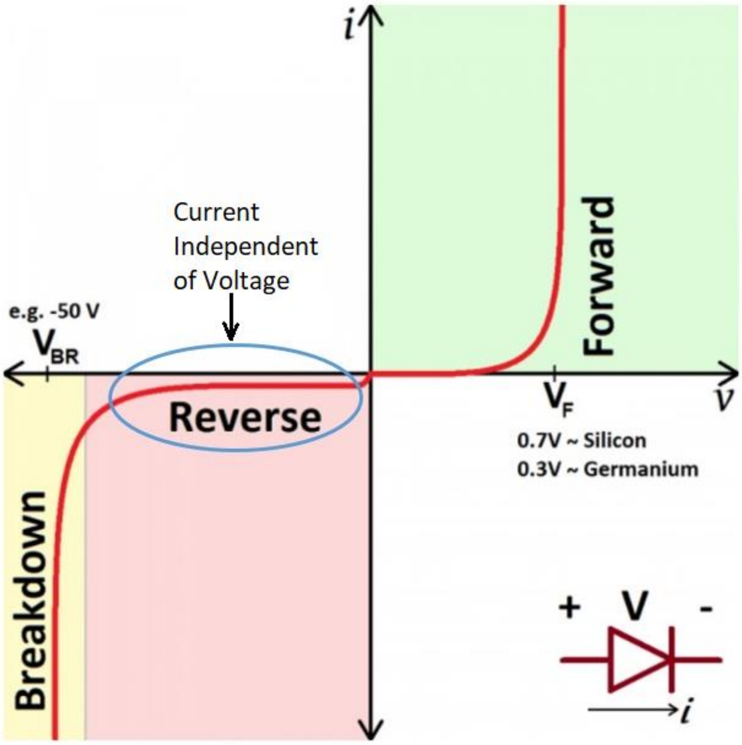


Fig. 1: Plot of the Diode Equation (1) (reproduced from [2]) (The scales in the positive and negative halves of the axes are unequal)

Now, we can clearly see from the above plot (Fig. 1) that the diode cannot be used to make a current source by operating it in the forward biased region. What about the reverse biased region? We see in the plot that in the reversed biased region, the current through the diode remains approximately constant, that is, it does not change with the voltage across the diode for upto a large value of this voltage (upto the point at which breakdown happens). So, this region of operation of the diode can be used to make a current source. The circuit connections for this operation are shown below (Fig. 2).

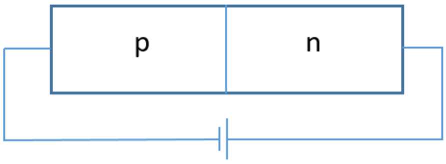


Fig. 2: Circuit Connections for Operating the Diode in the Reverse Biased Region by Doing Which It Can Be Used to Make a Current Source

We also note here that the current flowing through the diode in the reverse biased region is I_s, the reverse saturation current of the diode.

Now, how can we make this current source “voltage-controlled”? We note that the current flowing through the diode when operated with a reverse bias, the reverse saturation current, is due to the minority carriers, i.e., due to the minority electrons in the p-side of the diode flowing to its n-side and due to the minority holes in its n-side flowing to its p-side. So, we have to figure out how we can make the minority carriers of this diode voltage-controlled. Since we are talking about getting minority carriers such as getting electrons in the p-side, how about connecting an n-type semiconductor (from where we can get electrons) to the p-side of this diode (shown in Fig. 2) and applying a forward bias to this new diode as shown below (Fig. 3)?

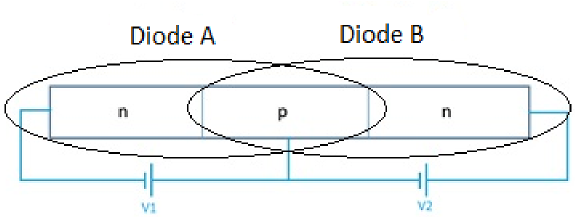


Fig. 3: A Scheme for Getting a Voltage-controlled Current Source

The current through the "n-p" diode on the left of Fig. 3 (diode A) — our new diode — depends on the voltage V1 across this diode according to equation (1). Now, this current brings electrons into the p-type section of the semiconductor in Fig. 3 from the n-type section on its left. Thus, the electrons in the p-type section of the semiconductor which constitute the minority carriers in this section are made to depend on the voltage V1 across the "n-p" diode on the left of Fig. 3 (diode A). Now, the reverse saturation current which flows through the reverse biased "p-n" diode on the right of Fig. 3 (diode B) is constituted majorly by the minority carriers (which are electrons) of the p-type section of the semiconductor. This reverse saturation current, thus, depends on the voltage V1 across the "n-p" diode on the left (diode A). We already know that this reverse saturation current does not depend on the voltage V2. Thus, the current flowing through the "p-n" diode on the right of Fig. 3 (diode B) is made to depend not on the voltage V2 across it, but on the voltage V1 across the "n-p" diode on the left (diode A). We thus have a voltage-controlled current source.

This is the fundamental idea of the bipolar junction transistor. The n-type section on the left of Fig. 3 is called the "emitter", the p-type section in the middle is called the "base", and the n-type section on the right is called the "collector".

From the above explanation, we can create a circuit model for the BJT. For the emitter-base junction on the left, we put an "n-p" diode, and for the collector-base junction on the right, we put a current-source (voltage-controlled) the arrow in which has the same direction as that of the "n-p" diode. This is shown below (Fig. 4).

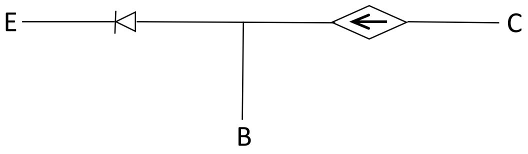


Fig. 4: Circuit Model for the BJT [3]

This is because as explained above, the emitter-base junction acts as a forward-biased p-n junction diode, and the reverse-biased collector-base junction acts as a current source (voltage-controlled, of course). By the by, this mode of operation of the transistor in which the emitter-base junction is forward-biased and the collector-base junction is reverse-biased is called the forward active mode of operation. The circuit model shown here is called the Ebers-Moll model of the BJT. What is a circuit model, by the by? We try to substitute the transistor having a complex I-V characteristics by a connection of circuit elements such that each element has a simpler I-V characteristics for ease of analysis of circuits. Now, let us see the currents and voltages in this model (Fig. 5).

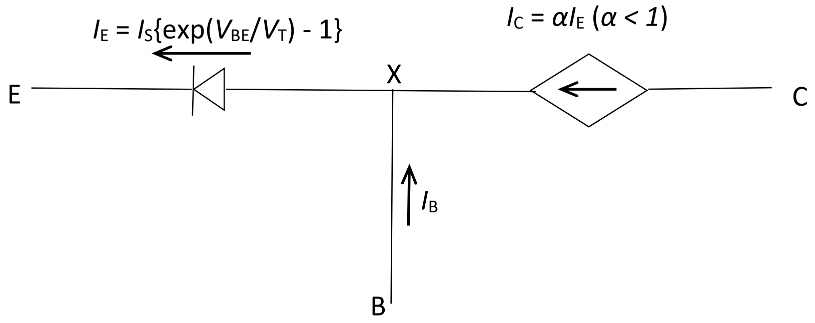

$I_E = I_S\{\exp(V_{BE}/V_T) - 1\}$
$I_C = \alpha I_E$ ($\alpha < 1$)
E
X
C
I_B
B

Fig. 5: Currents and Voltages in the Circuit Model of the BJT

We see that

$$I_E = I_S(e^{\frac{V_{BE}}{V_T}} - 1) \ldots (1)$$

following the I-V equation of the diode.

A part α of this emitter current comes from the collector and the remaining part comes from the base. The value of α is less than, but very close to 1 (≈ 0.98 or 0.99).

$$I_C = \alpha I_E = \alpha I_S(e^{\frac{V_{BE}}{V_T}} - 1) \ldots (2)$$

Applying Kirchhoff's Current Law (KCL) at node X, we have

$$I_B + I_C = I_E \ldots (3)$$

or, $I_B + \alpha I_E = I_E$

or, $I_B = (1 - \alpha)\, I_E \ldots (4)$

$\therefore\ \frac{I_C}{I_B} = \frac{\alpha I_E}{(1-\alpha)I_E}$

or, $\frac{I_C}{I_B} = \frac{\alpha}{1-\alpha}$

or, $\frac{I_C}{I_B} = \beta\ \left[where\ \beta = \frac{\alpha}{1-\alpha}\right]$

$\therefore I_C = \beta I_B \ldots (5)$

The value of β turns out to be 49 for $\alpha = 0.98$ and 100 for $\alpha = 0.99$.

Let's take a look at equation (2). We see that the current I_C depends only on V_{BE}. This equation helps us to appreciate that the BJT acts as a voltage-controlled current source.

Now, let $\alpha I_S = I_{S1}$

Then, $I_S = \frac{I_{S1}}{\alpha}$

We then have currents in the above circuit model as shown below (Fig. 6):

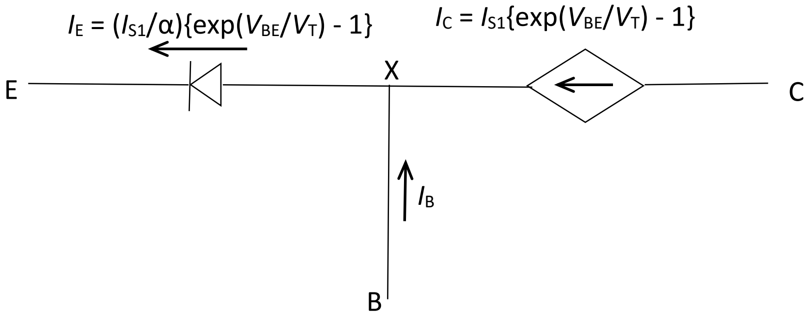

$I_E = (I_{S1}/\alpha)\{\exp(V_{BE}/V_T) - 1\}$
$I_C = I_{S1}\{\exp(V_{BE}/V_T) - 1\}$
E
X
C
I_B
B

Fig. 6: Usual Way of Showing the Currents in the Circuit Model of the BJT

This is the usual way of showing the currents and voltages in the circuit model of the BJT. There again, instead of I_{S1}, we write I_S. Then, the currents in the circuit model of the transistor appear as follows (Fig. 7):

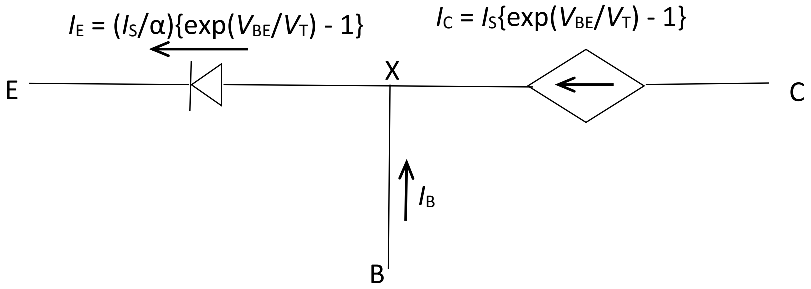

$I_E = (I_S/\alpha)\{\exp(V_{BE}/V_T) - 1\}$
$I_C = I_S\{\exp(V_{BE}/V_T) - 1\}$
E
X
C
I_B
B

Fig. 7: Usual Way of Showing the Currents in the Circuit Model of the BJT (Obtained by Replacing I_{S1} in Fig. 6 by I_S)

This is the circuit model of the transistor in the forward active mode of its operation. There is a reverse active mode of operation of the transistor [3]. In this mode, the collector-base junction is forward-biased and the emitter-base junction is reverse-biased. This mode of operation of the transistor is similar to the forward active mode of operation in that in this mode, the collector-base junction acts as a forward-biased diode and the emitter-base junction acts as a current source (voltage-controlled). But these two modes of operation are not exactly same. We differentiate between these two modes by writing for α α_F for the forward active mode and α_R for the reverse active mode, and also for I_S I_{SF} for the forward active mode and I_{SR} for the reverse active mode. We thus arrive at the following circuit model for the reverse active mode of operation (Fig. 8) [3].

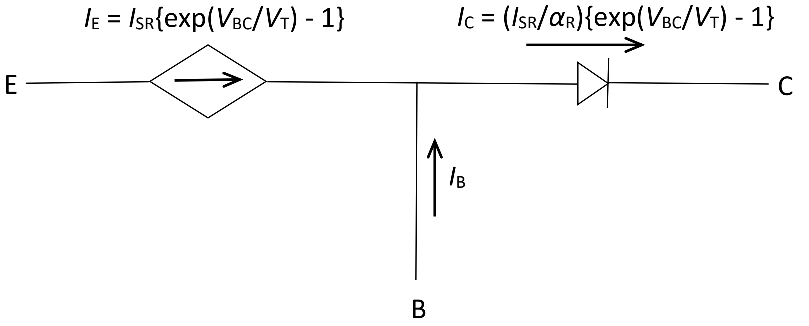
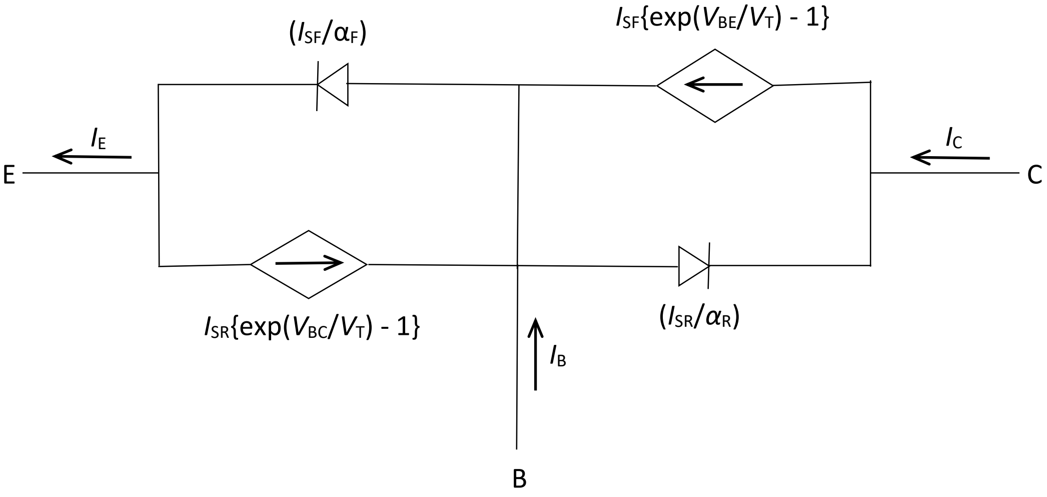

$I_E = I_{SR}\{\exp(V_{BC}/V_T) - 1\}$
$I_C = (I_{SR}/\alpha_R)\{\exp(V_{BC}/V_T) - 1\}$
E
C
I_B
B

Fig. 8: Circuit Model for the Reverse Active Mode of Operation of the BJT

The circuit models for the forward active and the reverse active modes of operation can be combined to give us the complete Ebers-Moll model of the BJT as shown below (Fig. 9) [3]:



(I_{SF}/α_F)
$I_{SF}\{\exp(V_{BE}/V_T) - 1\}$
I_E
E
I_C
C
$I_{SR}\{\exp(V_{BC}/V_T) - 1\}$
(I_{SR}/α_R)
I_B
B

Fig. 9: Complete Ebers-Moll Model of the BJT

In the above model, I_{SF}/α_F and I_{SR}/α_R are the reverse saturation currents of the two diodes. The complete Ebers-Moll model should be used to find out the equations relating the voltages and currents in the saturation region of operation of the transistor. We should remember that in the saturation region of operation of the transistor, both the emitter-base junction and the collector-base junction are forward biased.

As has been explained above, the following are the equations for the currents and voltages in the BJT in its forward active mode of operation:

$$I_C = I_S\left(e^{\frac{V_{BE}}{V_T}} - 1\right) \dots (6)$$

$$I_B = \frac{I_C}{\beta} = \frac{I_S}{\beta}\left(e^{\frac{V_{BE}}{V_T}} - 1\right) \dots (7)$$

$$I_E = I_C + I_B \dots (8)$$

Also, since $\beta = 100$ or 50, therefore, $I_B << I_C$.

Using these relationships, we arrive at the following circuit model of the transistor for its forward active mode of operation (Fig. 10) [4]:

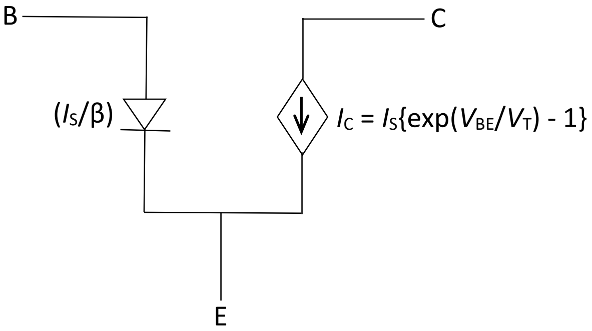


Fig. 10: A Second Circuit Model of the BJT (I_S/β is the Reverse Saturation Current of the Diode)

There is a way of drawing this circuit model without memorizing it. We first note that $I_B << I_C$, and hence, $I_E \approx I_C$. We, therefore, first join the collector (C) terminal to the emitter (E) terminal through a controlled current source (voltage-controlled) as follows (Fig. 11):

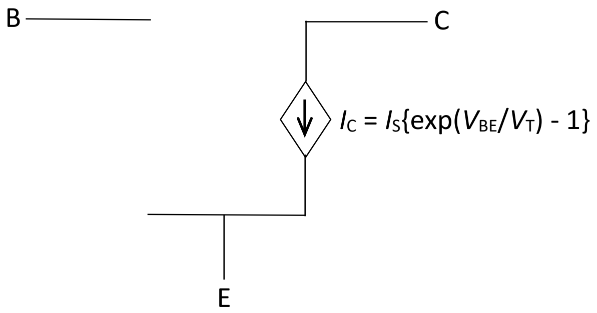


Fig. 11: First Step to Drawing the Second Circuit Model of the BJT

Now, we note that I_E is not exactly equal to I_C. The remaining part of I_E comes from the base (B) terminal. So, we join the terminal B to the terminal E through some circuit element through which the base current should be flowing. It is as follows (Fig. 12):

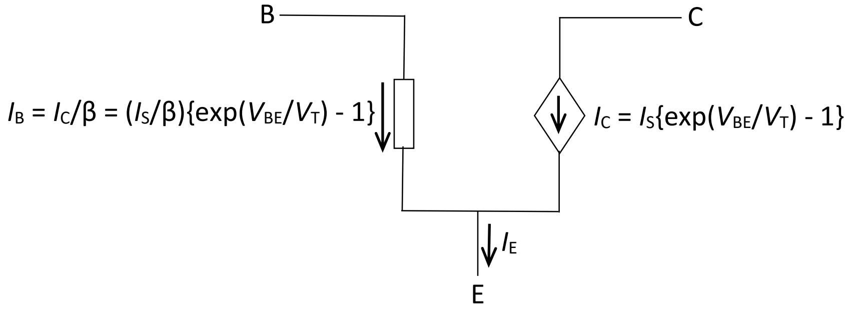

B
C
$I_B = I_C/\beta = (I_S/\beta)\{\exp(V_{BE}/V_T) - 1\}$
$I_C = I_S\{\exp(V_{BE}/V_T) - 1\}$
I_E
E

Fig. 12: Second Step to Drawing the Second Circuit Model of the BJT

We now have to figure out what the circuit element between the base (B) terminal and the emitter (E) terminal is. We note that since $I_B = I_C/\beta$, therefore, $I_B = \frac{I_S}{\beta}\left(e^{\frac{V_{BE}}{V_T}} - 1\right)$. Thus, the circuit element between B and E is such that the current through it is dependent on the voltage across it according to this equation. Hence, it must be a diode with a reverse saturation current of I_S/β. The resulting circuit model is as follows (Fig. 13):

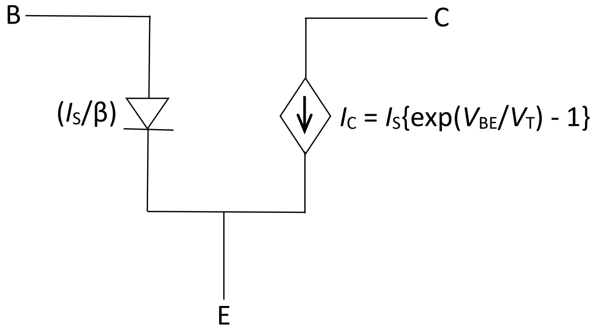


Fig. 13: Final Step to Drawing the Second Circuit Model of the BJT

We include the effect of the saturation region of operation of the BJT by adding a p-n junction diode between the p-type base and the n-type collector as shown below (Fig. 14) [4]:

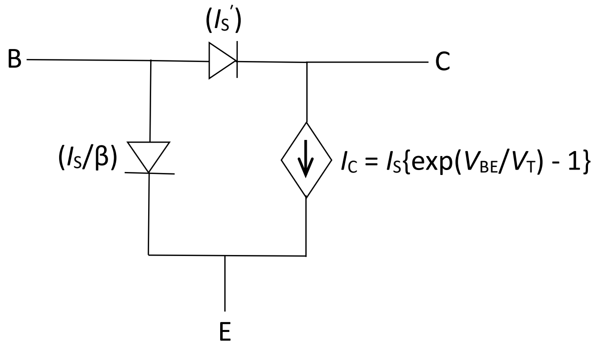


Fig. 14: The Second Circuit Model of the BJT Completed So As to Be Able to Take Into Account the Saturation Region of Operation

We should remember that in the saturation region of operation of the BJT, both the emitter-base junction and the collector-base junction are forward-biased, i.e., in our complete circuit model, both the base-emitter diode and the base-collector diode are forward-biased.

These circuit models define the BJT and enable us to figure out how we can use the BJT. These circuit models are called “large-signal models”. By the term “large-signal model”, we do not mean that the model can be used only when the applied signal is “large”. The model can be used also when the signal is “small”. We also note here that the model can be used both for dc as well as ac signals.

References:

1. Ben G. Streetman and Sanjay Kumar Banerjee, “Solid State Electronic Devices”, PHI Learning Private Limited, 6th Edition, 2009.
2. https://learn.sparkfun.com/tutorials/diodes/real-diode-characteristics
3. Adel S. Sedra and Kenneth C. Smith, “Microelectronic Circuits”, Oxford University Press, 5th Edition, 2004.
4. Behzad Razavi, “Fundamentals of Microelectronics”, Wiley, 2nd Edition, 2014.